DEPARTMENT OF NATIONAL DEFENCE
CANADA

OPERATIONAL RESEARCH DIVISION

DIRECTORATE OF OPERATIONAL RESEARCH (JOINT & LAND)

ORD PROJECT REPORT PR 9806

AN EVALUATION OF LOGISTICS PLANNING TOOLS

BY

Maj Doug Hoyle, JSORT 4
And
Michel Vigneault, DASOR 3

JULY 1998

OTTAWA, CANADA

National Defence Défense nationale

OPERATIONAL RESEARCH DIVISION

CATEGORIES OF PUBLICATION

ORD Reports are the most authoritative and most carefully considered publications of the DGOR scientific community. They normally embody the results of major research activities or are significant works of lasting value or provide a comprehensive view on major defence research initiatives. ORD Reports are approved personally by DGOR, and are subject to peer review.

ORD Project Reports record the analysis and results of studies conducted for specific sponsors. This Category is the main vehicle to report completed research to the sponsors and may also describe a significant milestone in ongoing work. They are approved by DGOR and are subject to peer review. They are released initially to sponsors and may, with sponsor approval, be released to other agencies having an interest in the material.

Directorate Research Notes are issued by directorates. They are intended to outline, develop or document proposals, ideas, analysis or models which do not warrant more formal publication. They may record development work done in support of sponsored projects which could be applied elsewhere in the future. As such they help serve as the corporate scientific memory of the directorates.

ORD Journal Reprints provide readily available copies of articles published with DGOR approval, by OR researchers in learned journals, open technical publications, proceedings, etc.

ORD Contractor Reports document research done under contract of DGOR agencies by industrial concerns, universities, consultants, other government departments or agencies, etc. The scientific content is the responsibility of the originator but has been reviewed by the scientific authority for the contract and approved for release by DGOR.

DEPARTMENT OF NATIONAL DEFENCE

CANADA

OPERATIONAL RESEARCH DIVISION

DIRECTORATE OF OPERATIONAL RESEARCH (JOINT & LAND)

ORD PROJECT REPORT PR 9806

AN EVALUATION OF LOGISTICS PLANNING TOOLS

by

Major Doug Hoyle, JSORT 4
And
Michel Vigneault, DASOR 3

Recommended by: _____
G. Lafond
DDR(J&L)

Approved by: _____
A.P. Bradfield
DGOR

ORD Project Reports present the considered results of project analyses to sponsors and interested agencies. They do not necessarily represent the official views of the Canadian Department of National Defence.

OTTAWA, ONTARIO

JULY 1998

102-02477/

ABSTRACT

In March 1997, J4 Log tasked LOGORT to conduct an evaluation of two software systems, PLANLOG and CANLOG Planner that were being considered as potential logistics planning tools for the Canadian Forces. A multi-criteria methodology for evaluating logistics planning software was developed. It proposed a list of criteria to be considered. Subject matter experts rated the criteria on their importance to logistics software selection. The two tools were assessed against the criteria during technical and user evaluations. During each of the evaluations, a subset of the criteria was evaluated using the judgements of evaluators, in order to produce satisfaction ratings. The satisfaction ratings were analysed to identify the strengths and weaknesses of the tools. Subsequently the satisfaction ratings were combined with the importance ratings for the criteria to generate overall performance results. The results of this analysis identified PLANLOG as a better logistics planning tool than CANLOG Planner.

RESUME

En mars 1997, J4 Log a assigné à ÉRO(L) la tâche d'évaluer deux outils planification logistique pour les Forces Canadiennes. Une méthode multi-critère pour évaluer les logiciels a été développée. Celle-ci propose une liste de critères à considérer. Des experts en planification logistique ont coté les critères selon leur importance pour le choix d'un logiciel de planification logistique. Les deux logiciels ont été évalués selon des critères tecnniques, puis fonctionnels. Durant chacune des évaluations, un sous-ensemble des critères a été considéré avec l'aide du jugement des évaluateurs, afin de produire des scores mesurant les forces et les faiblesses des logiciels. Ensuite, les scores de satisfaction ont été combinés aux scores d'importance pour calculer la performance générale. Les résultats de cette analyse suggèrent que PLANLOG est un meilleur outil de planification logistique que CANLOG Planner.

TABLE OF CONTENTS

PAGE

ABSTRACT i
TABLE OF CONTENTS ii
EXECUTIVE SUMMARY iii
BACKGROUND 1
OBJECTIVES 1
SCOPE 2
SOFTWARE TOOLS 2
PLANLOG 2
CANLOG PLANNER 3
EVALUATION METHODOLOGY 4
CRITERIA IMPORTANCE SURVEY 5
TECHNICAL EVALUATION 8
USER EVALUATIONS 12
WEIGHTED EVALUATION 16
CONCLUSIONS 17
RECOMMENDATIONS 18
REFERENCES 20
ANNEX A – EVALUATION CRITERIA A-1
ANNEX B – EVALUATION MEASUREMENT SYSTEM B-1
ANNEX C – PLANLOG USER EVALUATION C-1
ANNEX D – CANLOG PLANNER USER EVALUATION D-1

LIST OF TABLES

Table I: Importance Survey Results 7

LIST OF FIGURES

Figure 1: Satisfaction Scores for Benefit Criteria 9
Figure 2: Satisfaction Scores for Cost Criteria 10
Figure 3: Satisfaction Scores for Risk Criteria 11
Figure 4: Satisfaction Scores for Practicality Criteria 13
Figure 5: Satisfaction Scores for Employability Criteria 15
Figure 6: Weighted Evaluation Results 17

EXECUTIVE SUMMARY

In March 1997, DGOR was tasked to evaluate two logistics planning tools for possible employment in the Department of National Defence. The tools under consideration were CANLOG Planner, a Canadianized version of the US OPLOG Planner and PLANLOG, a Dutch logistics planning tool. The tools performed some common tasks but each had unique capabilities. Since both tools were available at limited cost, other tools were not considered initially. The challenge of comparing dissimilar tools along with the potential requirement to evaluate other tools in the future created a requirement for a reusable yet flexible evaluation methodology. Therefore, the project team developed a reusable methodology for logistics software evaluation, and employed it to evaluate the two logistics planning tools. The evaluation methodology and first part of the software evaluation project have been previously reported upon. This project report summarises the earlier work and reports the results of the remainder of the evaluation. Conclusions and recommendations from the entire body of work are provided.

The evaluation methodology that was developed employed a set of 28 evaluation criteria. The criteria were grouped by similarity of characteristics into five categories, Benefits, Costs, Risks, Practicality, and Employability. Not all criteria were suitable for evaluation in the same manner. Therefore technical experts evaluated three criteria groups and a user evaluation was employed for the remaining two criteria groups. Both sets of evaluators and users were given a standard questionnaire. The results of the two evaluations were analysed to identify the strengths and weaknesses of the tools.

It was determined that not all criteria were of equal importance to logistics software decision making. Therefore a criteria importance survey was conducted with major stakeholders in the logistics community to determine the relative importance of the evaluation criteria. After this step, the results of the criteria importance survey were used to create a criteria weight structure. The criteria weight structure was applied to the criteria satisfaction ratings to identify the weighted satisfaction score.

Representatives of each of the Environmental Staffs, J4 Log, CFCSC, and CFSAL completed the criteria importance survey. They rated each evaluation criteria on an importance scale of 1 (Totally Unimportant) to 9 (Totally Important). The results of the survey were combined to establish average importance scores for each criterion. These values ranged from 5.2 (Neither Important nor Unimportant) to 8.8 (Totally Important).

No criterion received an average score indicating that the criterion was unimportant. This was interpreted as a validation of the selected criteria. The importance scores were used to develop a weight structure. The weights were calculated so that no individual respondent could over-influence the relative importance of the criteria. The weights ranged from 4.3 for the most important criteria to 2.5 for the least important criteria. The weights were used as multipliers of evaluation results so that the scores of important evaluation criteria were emphasised over the scores of less important criteria.

The technical evaluation of the tools was conducted to identify the Costs, Benefits, and Risks associated with each of the tools. A technical evaluator rated the software on a scale of 1 (Totally Unsatisfactory) to 9 (Totally Satisfactory) for each of the criteria in the previously mentioned criteria groups. The results of this evaluation rated PLANLOG better than CANLOG Planner in two of the three criteria groups. The unweighted sum of the criteria scores for PLANLOG was 73.0 compared to CANLOG Planners score of 67.0. These results did not lead to any conclusion regarding the superiority of one tool over the other. They did however, indicate that a user evaluation of both tools was warranted.

Logistics practitioners conducted the user evaluations of CANLOG Planner and PLANLOG after they had become familiar with the tools. Employing the same satisfaction scale as was used for the technical evaluation, they rated the criteria from the remaining two criteria groups. PLANLOG received higher satisfaction scores than CANLOG Planner in both criteria groups. The combined user evaluation satisfaction scores for PLANLOG were 84.8 compared to CANLOG Planners 65.5.

Several global performance measures were employed. The sum of the satisfaction scores for all evaluation criteria were compared. PLANLOG scored seventeen percent higher than CANLOG Planner. As a final performance measure the weight structure derived from the importance survey results were used to multiply criteria satisfaction scores. The effect of this calculation was to emphasise the satisfaction ratings of the important criteria over less important criteria. The weight structure for this analysis increased the most important criteria scores by 60 % over the least important criteria. The weighted satisfaction ratings confirmed that PLANLOG was a better tool than CANLOG Planner. PLANLOG had higher weighted scores in four of the five criteria categories and PLANLOG scored eighteen percent higher than the CANLOG Planner. This increase in relative scores when the importance of criteria was considered strengthened the initial finding of the evaluation.

The evaluation process identified the strengths and weaknesses of each tool and quantified the importance of these characteristics to the software selection process. It enabled the project team to conclude that PLANLOG is a better choice for employment as a joint staff logistics planning tool than CANLOG Planner. Therefore, it is recommended that PLANLOG be acquired from TNO Laboratories after the necessary changes identified during the evaluation are implemented.

This project proved that this evaluation methodology is an effective way to evaluated logistics planning software. The methodology can easily be adapted to evaluate other types of software. Although some evaluation criteria would need to be changed, the general approach to software evaluation would remain the same. It is recommended that this methodology be employed when other tools are considered.

AN EVALUATION OF LOGISTICS PLANNING TOOLS

BACKGROUND

1. In March 1997, LOGORT was tasked by J4 Log to conduct an evaluation of two logistics planning tools (Ref. 1, 2). This Software evaluation project was a result of several independent events during the fall of 1996, which brought to the attention of logistics staff at National Defence Headquarters a requirement to evaluate logistics planning tools. The first event was the development of CANLOG Planner, a version of a US logistics-planning tool OPLOG Planner adapted for Canadian use by 1 Canadian Division Headquarters at Kingston. Secondly, the Netherlands research organisation, TNO, offered Canada their tool, PLANLOG as part of a bilateral collaboration agreement. In addition, several other logistics planning software initiatives were identified. The duplication and lack of co-ordination between these software initiatives became a problem that needed to be addressed. The Software Evaluation Project was started to define a reusable evaluation methodology to deal with logistics planning tools for the Canadian Forces. The methodology was applied to the two tools cited above. In February 1998 a research note was published which outlined the evaluation methodology and the results of the first part of the evaluation process.

OBJECTIVES

2. Logistics planning encompasses a wide variety of functions such as consumption planning, distribution planning, and maintenance planning. Not surprisingly, tools that perform logistics planning may be very different from one another. Indeed the two logistics planning tools that were under consideration in Canada worked very differently. This situation made comparison of these logistics planning tools very difficult. The solution to this dilemma required the identification of the strengths and weaknesses of each tool as well as how important these characteristics were to the logistics planning process. The overall objective of this project was *to determine which tool was most suitable for logistics-planning in Canada.* This broad objective implied that all reasonable candidate software packages would be evaluated in an objective and comprehensive manner. Initially, only two software tools were considered because they

met the sponsor's criteria of low cost and immediate availability. However, to satisfy the primary objective, it was understood that we might have to consider a number of other logistics software tools sometime in the future. For this reason, we adopted a sub-objective to *develop a baseline evaluation plan for future software systems.*

SCOPE

3. The evaluation methodology employed in this project was conceived with the aim of its reuse in evaluations of other logistics software tools. The mature version of this methodology is to be used by logistics experts without operational research support. The methodology has been explained in a previously published research note (Ref. 3). The note describes the evaluation methodology and the results of the technical evaluation of the CANLOG Planner and PLANLOG tools. This project report summarises the previously reported work and presents the results of the remainder of the evaluation project. Conclusions and recommendations based upon the entire body of work for this project are presented.

SOFTWARE TOOLS

PLANLOG

4. PLANLOG is a logistics planning tool developed by the Dutch research organisation TNO for the Royal Netherlands Army. This program was built in Delphi and uses the Paradox database access software. PLANLOG came to our attention through a review of potential candidates for a Dutch-Canadian collaboration program. It was one of the tools that could potentially fill a Canadian requirement in logistics software.

5. PLANLOG estimates consumption of food (Class I), fuel (Class III) and ammunition (Class V) for operational units and for the logistics units supplying them. The units and routes can be drawn over a bitmap representation of the theatre of operation. PLANLOG rolls up consumption from staff data, applying intensity factors depending upon the activity, the type of terrain and the weather conditions. PLANLOG also allocates transportation resources using a linear optimisation solver. Although the original Dutch version allocated transportation resources by stock class, TNO produced, at our request, a version that allocates transportation by supply subclass. The subclasses were created to ensure compatibility of goods for transportation. This change made the number of variables in the optimisation model too large for a solver to handle in a reasonable time period. Therefore the problem was relaxed to searching for an optimal non-integer allocation of transportation means, then rounding up the solution to an integer allocation. Therefore even though the solutions found in the tested version were

very good, they were not necessarily optimal. For moderately complex scenarios, the solver was still too slow; TNO then improved the tool by setting a variable time limit on the solver. With this version, we could perform exploratory analysis with a short time limit, and also find solutions closer to optimal by setting a longer time limit. This design provided the best balance between solution quality and computational efficiency.

6. Since the completion of the PLANLOG evaluation, TNO has completed major enhancements to their tool's functionality. Among the improvements are the ability to plan for all classes of supply, the ability to track stock holdings over time and an enhanced reporting ability. They have also included medical and maintenance planning modules. These modules have extensive functionality that had not been investigated when this report was completed. Other recent developments concerning PLANLOG include the inclusion of a commercial optimizer in PLANLOG. This means that Canada will have to pay licensing fees of approximately $100.00 for each copy of PLANLOG it chooses distribute. Finally the Royal Netherlands Army has selected PLANLOG for inclusion in a suite of planning tools they will be fielding in the future. This development reduced the risk that the Dutch could abandon PLANLOG in the near future.

CANLOG PLANNER

7. CANLOG Planner is a staff data planner that performs estimates of the consumption of most classes of supply. It has been adapted from the American OPLOG Planner. CANLOG Planner employs Canadian staff data (when available) and has a customized user interface. The rest of the tool is identical to OPLOG Planner. The program is the property of the DoD and it maintains all ownership over the product. It also controls the maintenance and development of this relatively mature tool. Headquarters 1 Canadian Division (HQ 1 Can Div) has negotiated permission for DND to use a version of it. The tool was released to the army as part of the Electronic Battle Box planning tool suite in the fall of 1997.

8. Starting with a combination of user defined data and staff data, CANLOG Planner generates three types of estimates: AMMUNITION expenditures, POL consumption and OTHER CLASSES OF SUPPLY consumption. Ammunition consumption is calculated based upon the quantity of each type of weapon in the units employed and the intensity of the operation. POL consumption is calculated based upon the number of each type of vehicle, generator or heater and their expected employment. The OTHER CLASSES OF SUPPLY consumption report is calculated based upon the number of personnel and the conditions under which they are deployed. The estimates provided by CANLOG Planner are high level estimates, which are useful for planning purposes only. The Ammunition estimate will identify the weight and the cubic dimensions of the expected ammunition

consumption. This information is reported by ammunition type for each unit employed. The POL estimate provides an estimate, in litres and tons, of the consumption of each type of POL for each unit employed. The OTHER CLASSES OF SUPPLY report provides an estimate, in tons, of each remaining class of supply except spare parts (Class VII).

9. Since the evaluation of CANLOG Planner has been completed, the Department of Army Doctrine has contracted a firm to upgrade the tool. CANLOG Planner is being re-engineered, at considerable cost, to increase its user friendliness and make data maintenance easier. The new version will be included in the second release of the CD-ROM based electronic tool: "Electronic Battle Box". This release is schedule for the fall of 1998 (Ref 4).

EVALUATION METHODOLOGY

10. To assess software in a comprehensive manner, evaluation criteria are required. Our evaluation criteria were derived from the expert opinions of logistics practitioners and operational research analysts. A set of 28 evaluation criteria was developed. The criteria were identified and defined in reference 3. They have been included for easy reference in Annex A. The criteria can be grouped into five categories: Benefits, Costs, Risks, Practicality, and Employability. Due to the diversity of the criteria groups, the same group of experts could not adequately evaluate all of them. Therefore a two step evaluation strategy was devised. Technical experts conducted an evaluation of the first three categories of criteria. Logistics planning experts performed a user evaluation of the tools to report upon the final two evaluation categories.

11. An evaluation scale was developed to measure software performance. During the technical evaluation or the user trial, each tool was assigned a satisfaction rating between 1 (Totally Unsatisfactory) and 9 (Totally Satisfactory) for each evaluation criterion. These ratings formed a basis for comparison of the different software tools. In many cases, criterion satisfaction levels were defined by setting a baseline level at one end of the satisfaction spectrum and adjusting other levels relatively to that baseline. In most cases, objective definitions were not possible or suitable. A detailed explanation of the satisfaction rating system has been enclosed at Annex B.

12. All evaluation criteria were not equally important to the decision making process. Therefore it was necessary to identify the relative importance of the evaluation criteria so that the results of the technical and user evaluations could be interpreted correctly. Criteria importance ratings were acquired from a survey of the primary stakeholders in the logistics community.

13. Once the criteria importance survey, technical evaluation, and user evaluation were completed, an analysis of the results was conducted. Initially, the evaluation criteria satisfaction scores for the tools were compared to identify obvious strengths and weaknesses. Subsequently the criteria importance ratings were employed as a weight structure to emphasise the criteria satisfaction scores based upon their relative importance. The sum of the weighted and unweighted criteria scores were used as overall ratings for the two tools. Conclusions and recommendations were derived from this information.

CRITERIA IMPORTANCE SURVEY

14. The criteria importance survey was conducted to determine the relative importance of the evaluation criteria from the opinion of logistics stakeholders. The survey required the respondents to rate all evaluation criteria on a scale of 1 (Totally Unimportant) to 9 (Totally Important). Representatives from each of the environmental chiefs of staff (ECS), CFSAL, CFCSC, and J4 Log provided input. Annex B contains an explanation of the importance ratings that were employed.

15. The responses from the importance survey were used to create a criteria weight structure. The weights were calculated in the following manner. For a set of respondents of size m, let

$$Weight_i = \sum_{j=1}^{m}((Rating_{ij} \div Total\ Rates_j) \div m) \times 100 \qquad \forall i$$

Where $Weight_i$ is the calculated weight for criteria i, $Rating_{ij}$ is the rating for criteria i by respondent j, and $Total\ Rates_j$ is the sum of all criteria scores for respondent j. This method of calculating weights ensures that individual respondents do not have excessive impact on the weight calculation due to consistent over-scoring or under-scoring. These weights can be used as a multiplier of evaluation criteria scores to obtain a measure of the overall performance of a software tool

16. Table 1 contains the results of the survey. The table contains a column for the responses of each of the ECS, and one for the Joint Staff. The Land Staff submitted three sets of importance scores, which were provided by experts in different support disciplines (Supply, Transportation, and Maintenance). The Maritime and Air staffs provided one response each. The J Staff column contains the average of seven submissions from J4

Log, CFCSC and CFSAL. These respondents were grouped together since they all have a joint force outlook on logistics issues. The mean column contains the average of the four rating columns. This represents the average importance rating for each evaluation criteria. These results were examined for any obvious sign of bias. Generally, the responses were normally distributed about the mean value.

17. Table 1 also contains a column of calculated weights. The results of the importance survey revealed some important information about logistics software evaluations. The maximum importance score assigned to a criterion was 8.8 (Relevance of Output) and the minimum assigned score was 5.19 (Cost of Acquisition and Cost of O&M). Since no criterion were scored below 5 (Neither Important nor Unimportant), it was concluded that all criterion were worthy of consideration in the evaluation process. The cost criteria group received the lowest importance ratings of all criteria groups. Respondents indicated in their written comments that they were more concerned with acquiring effective planning tools that were easy to use than with getting the lowest cost solution. Most stakeholders demonstrated similar logistics software priorities regardless of their current place of employment. There did not appear to be significantly partisan replies

18. The weight values that were calculated for the evaluation criteria ranged from 4.21 (Relevance of Output) to 2.52 (Cost of Acquisition and Cost of O&M). Since 28 criteria were employed the average weight for each criteria would have been 3.6. These weights, when used as multipliers for criteria scores, increased the impact of the most important criteria by 60 % over the least important criterion. Although these weights appeared to be reasonable to the project team, the sponsor had the opportunity to adjust them if they believed that they were inappropriate. For example, if a very limited budget had been available for acquisition and maintenance of a tool, it would have been reasonable to increase the weights of the cost criteria. For the purpose of this software evaluation the weights were not adjusted.

TABLE 1

IMPORTANCE SURVEY RESULTS

CRITERIA	RATINGS				Mean Rating	Weight
	Land	Air	Maritime	JSTAFT		
Functionality						
User Friendliness	7.33	9.00	9.00	8.13	8.37	4.07
Learning Curve	8.67	7.00	9.00	7.80	8.12	3.97
Relevance Of Output	8.33	9.00	9.00	8.87	8.80	4.29
User-Confidence In Output	7.67	9.00	9.00	8.93	8.65	4.21
Data Sharing	7.00	7.00	8.00	8.67	7.67	3.74
Employability						
Army	5.00	8.00	9.00	6.89	7.22	3.50
AirForce	5.00	8.00	9.00	6.89	7.22	3.50
Maritime Forces	5.00	8.00	9.00	6.89	7.22	3.50
Joint Forces	7.33	8.00	7.00	8.67	7.75	3.78
OOTW	8.33	8.00	7.00	7.58	7.73	3.78
Crisis/War	8.00	8.00	9.00	8.56	8.39	4.09
What-If Scenario Analysis	7.00	8.00	9.00	8.89	8.22	4.01
Pre-Deployment	9.00	8.00	8.00	8.89	8.47	4.14
Deployment	8.33	8.00	8.00	8.56	8.22	4.02
Employment	8.00	8.00	8.00	8.89	8.22	4.01
Benefits						
Data Maintainability	8.00	7.00	8.00	8.67	7.92	3.87
Growth Potential	7.67	8.00	7.00	7.08	7.44	3.63
Completeness	8.33	7.00	7.00	7.08	7.35	3.60
Quality Of Results	8.67	7.00	7.00	8.08	7.69	3.76
Doctrinal Fit	6.33	5.00	6.00	6.67	6.00	2.94
Supportability	7.33	3.00	8.00	7.75	6.52	3.20
Costs						
Development	4.00	8.00	5.00	4.08	5.27	2.55
Acquisition	4.67	7.00	5.00	4.08	5.19	2.52
Training	4.00	8.00	5.00	4.08	5.27	2.55
O&M	4.67	7.00	5.00	4.08	5.19	2.52
Risks						
Senior Management Support	6.67	7.00	7.00	5.44	6.53	3.18
User Acceptance	7.67	9.00	6.00	7.22	7.47	3.64
Obsolescence	7.33	9.00	6.00	5.78	7.03	3.42

TECHNICAL EVALUATION

19. The technical evaluation team evaluated their level of satisfaction with the technical aspects of the software. Since one member of the technical evaluation team had helped develop the Canadian version of PLANLOG, he could not objectively evaluate that tool. Therefore, each member of the technical evaluation team was assigned a tool with which they were not too familiar. These initial evaluations were performed in isolation but they were followed up by a team review to discuss findings and to ensure that consistent assumptions were employed.

20. The separate technical evaluations of the software systems revealed some of the strengths and weaknesses for each system. The following comments provide a very general comparison of the results observed; specific findings for each software tool were provided in reference 3. The results of the technical evaluation are presented by criteria group.

21. Six benefit criteria were considered in the technical evaluation. The results are presented in Figure 1. The largest difference between the benefit criteria scores of the two systems was for **growth potential**. Evaluators noted PLANLOG's flexible programming technology and the chance or shared (CA/NL) control of forecasted growth, whereas CANLOG Planner would required a complete rewrite in order to expand its capabilities, and DND had little control over its future. Both systems obtained varying degrees of dissatisfaction for **data maintainability**, completeness and supportability. Regarding the data maintainability criteria, evaluators noted the lack of mechanisms for generating and verifying the staff data used by both systems. This was a weakness within DND rather than a weakness with either tool. While only U.S. DoD programmers could perform data updates to CANLOG Planner, PLANLOG had a user-friendly data edit program. For the **completeness** criteria, the limited scope of PLANLOG and the very limited scope of CANLOG Planner were noted. With respect to **supportability**, evaluators noted the lack of Canadian control over the continued support of each tool in their originating countries. The evaluators were ambivalent about the **quality of results** criteria for both systems. Results appeared reasonable, but in the case of CANLOG Planner, the presentation lacked detail and the reporting format was inflexible; PLANLOG's method of allocating transportation resources gave non-optimal results and it's reporting capabilities were rudimentary. Both systems obtained relatively high scores for **doctrinal fit**, each appearing to accommodate Canadian doctrine adequately within their limited scope of capabilities. PLANLOG received an overall score of 31 while CANLOG Planner received a total score of 24.

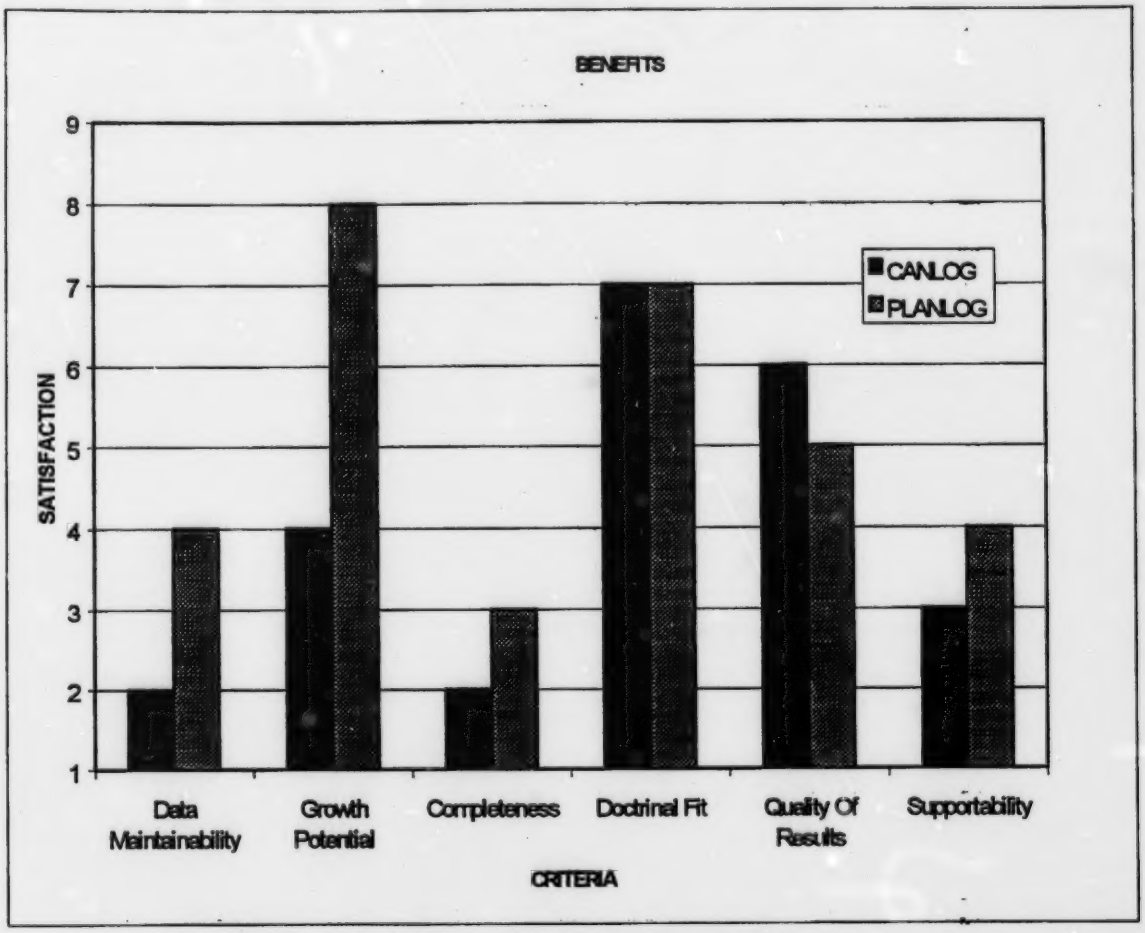

Figure 1: Satisfaction Scores for Benefit Criteria

22. The results of the evaluation of cost criteria are presented in Figure 2. All cost criteria rated fairly high for both software tools because of the excellent terms the tools were being received under. The **acquisition costs** for both tools were minimal because of the co-operation agreements with the nations that initially developed them. The U.S. had given CANLOG Planner to Canada while the Dutch had only asked for scientific collaboration in return for PLANLOG. **Development costs** were also generally satisfactory, since all CANLOG Planner developments costs and most PLANLOG development costs had already been spent. The evaluators were slightly more ambivalent regarding the costs of training and O&M. Both tools benefited from the potential sharing of these costs. However, CANLOG Planner was expected to incur slightly higher **training costs** because it required greater knowledge and skill to use it. With regards to **O&M costs**, Canada could have been expected to participate in the further development of PLANLOG while the U.S. was solely responsible for CANLOG Planner upgrades and changes to the current functionality were not anticipated at the time of the evaluation.

CANLOG Planner received a slightly higher total satisfaction score for cost criteria of 29.0 compared to the PLANLOG score of 26.0.

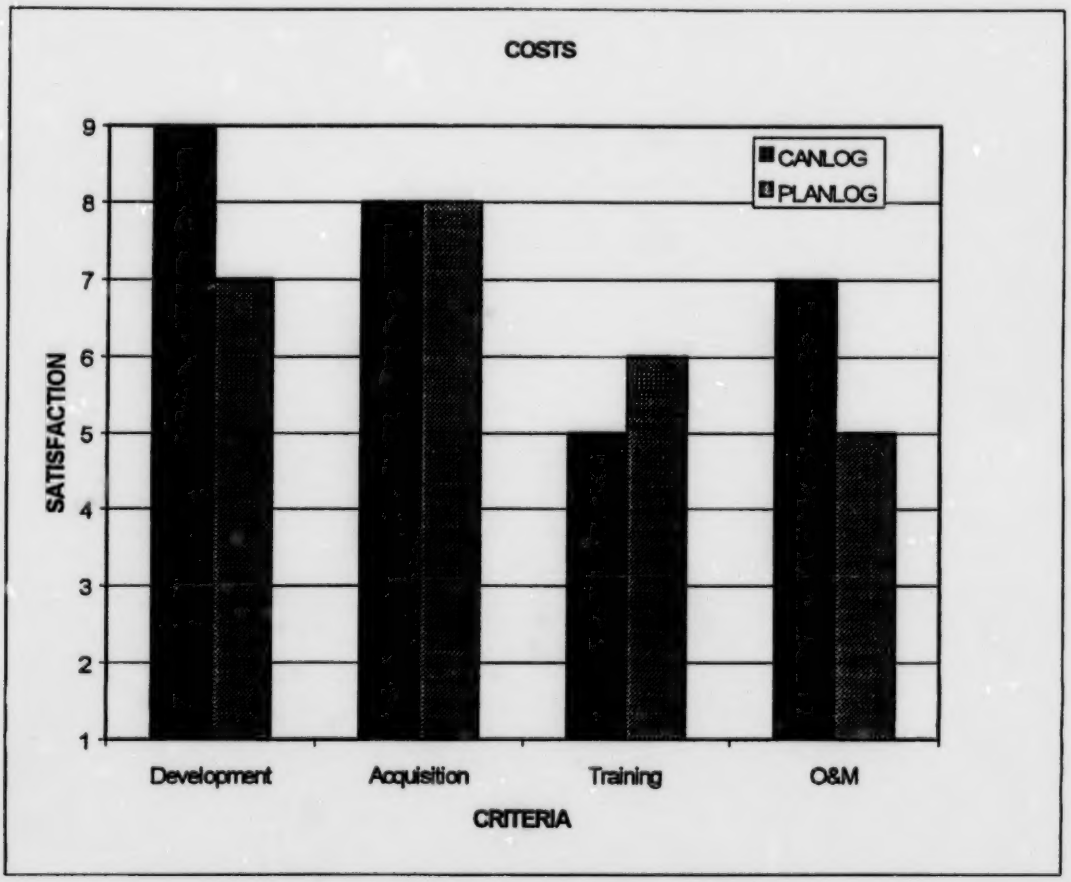

Figure 2: Satisfaction Scores for Cost Criteria

23. The results of the evaluation of risk criteria are contained in Figure 3. In the case of the **risk of senior management support**, CANLOG Planner obtained a better score because it is already available for operational use through the Electronic Battlebox and it had the support of senior management in US DoD. However, PLANLOG obtained a better score for **risk of user acceptance** because of its friendly and intuitive user interface, compared to CANLOG Planners older design/interface and its lack of flexibility. PLANLOG also obtained a better score for **risk of obsolescence** because it was developed using a modern programming language. Also, no other known tool duplicates PLANLOG's transport allocation capability. CANLOG Planner's total risk satisfaction score was 14.0 compared to PLANLOG's score of 16.0.

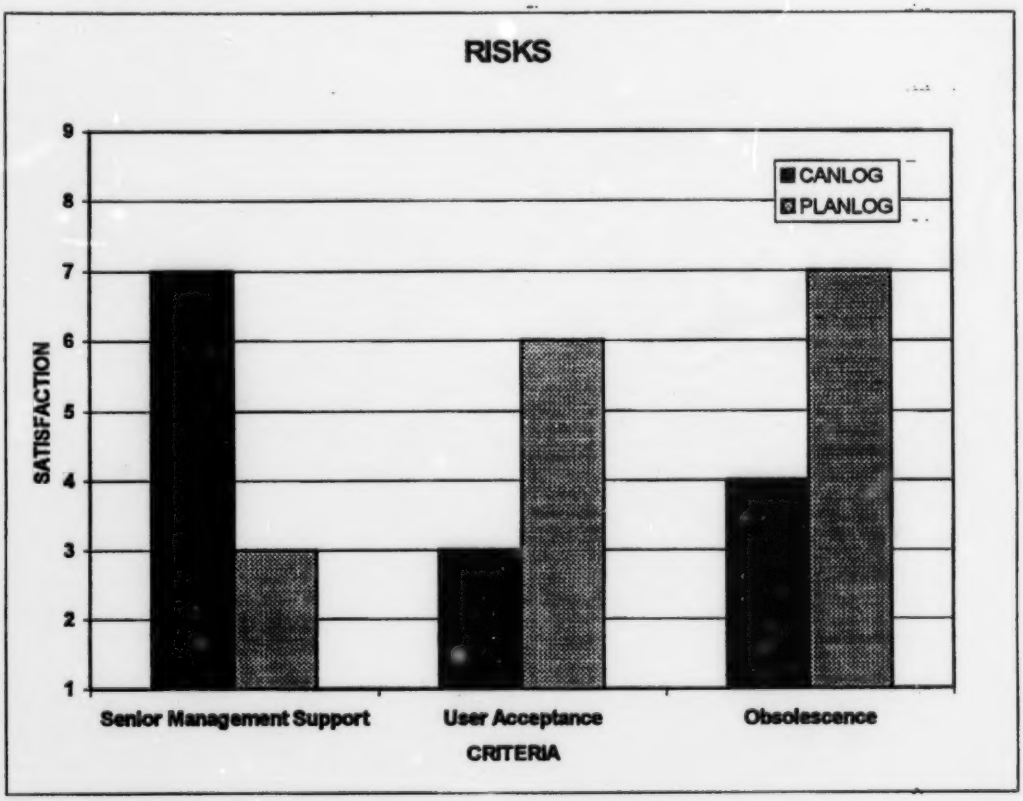

Figure 3: Satisfaction Scores for Risk Criteria

24. As a result of the technical evaluation it was concluded that each tool had unique strengths and weaknesses. CANLOG Planner was favoured in 4 criteria, PLANLOG was favoured in seven criteria and there were two criteria with tie scores. PLANLOG received a total score of 73.0 compared to CANLOG Planner's score of 67.0. Without prioritising evaluation criteria these results indicated that PLANLOG was technically stronger than CANLOG Planner was.

25. The results of the technical evaluation were analysed prior to the next stage of the evaluation. Due to the amount of time and effort required to conduct a user evaluation on software, the necessity of this step for both tools had to be confirmed to the sponsor. Although the technical evaluation results could not provided enough evidence to determine which tool was better than the other, they could have identified critical technical weaknesses that would have disqualified a tool from further consideration. In this case no critical technical weaknesses were identified and a user evaluation of both tools was justified. Therefore a user evaluation of both tools was conducted.

26. Since the technical evaluation was conducted, the status of the tools has changed. Some of these changes, which were outlined in paragraphs six and nine of this report, would have effected some of the satisfaction scores if the technical evaluation were conducted again. The net effect of these developments however, would not have changed the conclusions drawn from the evaluation. Therefore the impact of these recent developments are not considered elsewhere in this report.

USER EVALUATION

27. The user evaluation examined the functional characteristics of the logistics planning tools. Users worked with the software prior to completing a questionnaire about their experiences and impressions. Employing the previously explained satisfaction scale, they made judgements concerning the functional criteria. Evaluators were also encouraged to provide written comments where appropriate. Two kinds of functional criteria were evaluated: practicality criteria and employability criteria. A general explanation of the criteria and guidelines for the corresponding satisfaction levels is provided in Annex A. A formal user evaluation was conduct for PLANLOG. With sponsor support, nine Logistics experts from J4 Log, CFSAL and CLS were identified as participates in the evaluation. Evaluators were provided with an evaluation package. This package contained; a copy of the software, instructions on how to install the tool on their computers, a users manual and an evaluation instruction. Attached, as Annex C is the instructions the evaluators received and the table they employed to record their conclusions. The PLANLOG user evaluation went extremely well.

28. Logistics experts possessing prior experience with the CANLOG Planner were selected for the second user evaluation. Although it was assumed that many logistics planners would have prior experience with CANLOG Planner due to its inclusion in the Electronic Battle Box, this was not the case. After extensive investigation only six experienced users could be found to complete the CANLOG Planner user evaluation. In the future, a user evaluation similar to what was conducted for PLANLOG should be employed for all tools to ensure sufficient feedback on the tools. Attached, as Annex D is the evaluation instructions CANLOG Planner evaluators received

29. The results of the user evaluations were compiled and analysed. The satisfaction ratings provided by the respondents were grouped by tool and averaged to determine average satisfaction ratings. The practicality criteria group measured how well the logistics planning tools conformed to the users' practical needs. The five criteria in the group were user friendliness, learning curve, relevance of output, user-confidence in output and data sharing. The results of the practicality criteria evaluation are contained in Figure 4. This criteria group generated the greatest difference in satisfaction ratings

between the two tools. CANLOG Planner achieved very poor satisfaction ratings in this criteria group with all criteria receiving satisfaction scores below five. These consistently low scores in the practicality criterion suggested that the tool would require significant improvements for it to be successfully implemented. Textual comments indicated that these low scores were the result of user frustration with the design and age of the tool. The user interface of CANLOG Planner was not found to be intuitive, flexible, or user-friendly as indicated by the unsatisfactory/very unsatisfactory scores in the User Friendliness. Users experienced difficulty learning the necessary procedures to generate output with CANLOG Planner resulting in very low Learning Curve ratings. User Confidence in Output was rated mildly unsatisfactory because the tool gave a "black box" impression to the user. Users were unable to understand how the inputs to the tool affected the results that are generated.

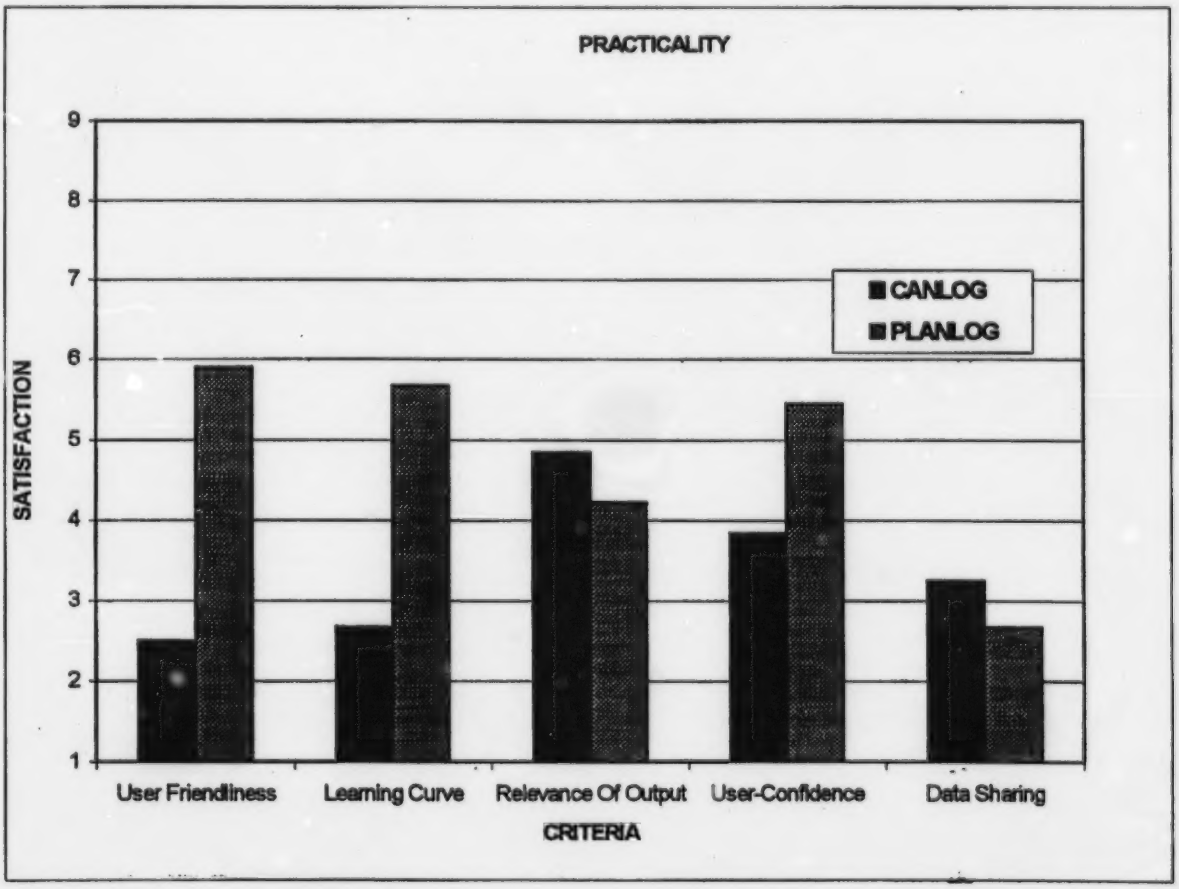

Figure 4: Satisfaction Scores for Practicality Criteria

30. PLANLOG received acceptable scores in most of practicality criteria however **data sharing** and **relevance of output** criteria scores indicated areas that require attention. User comments included suggestions on how to improve the practicality of PLANLOG. They expressed confidence that minor changes in the tool could result in significant improvements. The lack of high scores for either tool indicated that users had high expectations for the practicality of their logistics planning tools. With increased computer literacy in the CF these expectations are likely to increase and place greater demands on future tools. PLANLOG outscored CANLOG Planner in this criteria group by a total unweighted score of 23.9 to 17.0.

31. The employability criteria provided a way for the users to determine the limitations of the logistics planning tools. Three types of employability criteria were considered: the service (army, air force, maritime, joint forces), the level of conflict (OOTW, crisis/war) and the planning stage (what-if scenario analysis, pre-deployment, deployment, employment). With the exception of one criterion, users expressed greater satisfaction with PLANLOG than with CANLOG planner. This overwhelming result suggested a greater flexibility of employment for PLANLOG over CANLOG Planner. This point should be an important factor when tool selection decisions are made. PLANLOG was rated above five in eight of ten criteria where CANLOG Planner was rated at or below five on seven criteria. The most significant difference between the scores of the two tools was in the **Army** and **Employment** criteria. In both cases PLANLOG received satisfactory scores and CANLOG Planner received mildly unsatisfactory scores. Since both tools were originally developed for land forces, the large difference in **Army** employability scores became more significant. This fact also helps to explain why both tools scored poorly in **Maritime** criteria. Maritime logistics planning is conducted very differently than land logistics planning. Both tools will require some adaptation for them to be useful in the maritime environment. In the employability criteria group PLANLOG clearly outscored CANLOG Planner. PLANLOG received a total unweighted satisfaction score of 60.9 while CANLOG Planner only scored 48.5.

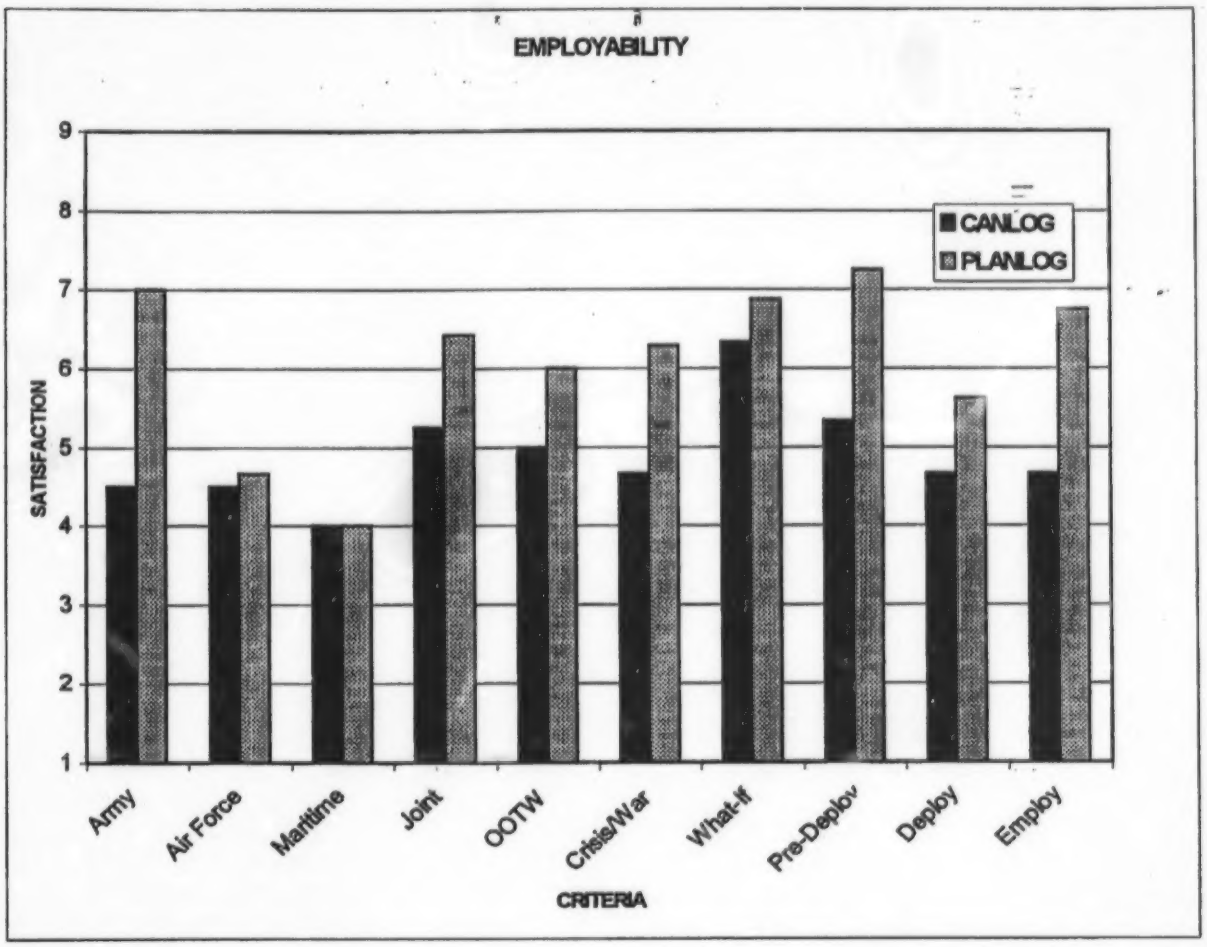

Figure 5: Satisfaction Scores for Employability Criteria

32. The user evaluation provided useful insights into the strengths and weaknesses of the tools. PLANLOG received a total functional criteria satisfaction score of 84.8 compared to the CANLOG Planner score of 65.5. The superiority of PLANLOG over CANLOG planner in both groups of functionality criteria has provided a good indication of the potential for success for these two tools in the CF. The textual comments indicated that there was a lot of potential for PLANLOG to be improved upon. Comments regarding CANLOG Planner indicated that the difficult user interface and complexity of employment significantly reduce the potential usefulness of the tool.

33. The satisfaction scores for all criteria were added to create a total satisfaction score for each tool. PLANLOG earned a score of 157.8 and CANLOG Planner earned a score of 132.5. This score represented sixteen percent higher satisfaction rating for PLANLOG over CANLOG Planner. These scores, although informative, did not provide conclusive evidence regarding the relative merits of the two tools. Without considering the relative importance of the evaluation criteria and the scores achieve in those criteria, incorrect conclusions could have been drawn.

WEIGHTED EVALUATION

34. The weights derived from the criteria importance survey were used as multipliers for the criteria satisfaction scores. The effect of the multipliers was to emphasise the satisfaction scores of the criteria with the higher weights. Weighted scores were calculated in the following manner:

$$Weighted\ Score_i = Weight_i \times Average\ Score_i \qquad \forall i$$

Where $Weighted\ Score_i$ was the weighted satisfaction score for criteria *i*, $Weight_i$ was the weight for criteria *i* derived from criteria importance survey, and $Average\ Score_i$ was the average of the satisfaction ratings for criteria *i*.

35. Figure 6 displays the sum of the weighted scores for each criteria group. This table illustrates how the two tools compared in each criteria group when criteria scores and criteria importance were considered. One cannot compare different criteria group because there are unequal numbers of criterion contributing to the total score in each group. PLANLOG achieved higher weighted scores than CANLOG Planner in four of the five criteria groups. The practicality criteria group exhibited the largest percent difference in weighted scores. This was consistent with the unweighted satisfaction scores that have been previously discussed. The total of all weighted criteria scores provided a useful performance measure. PLANLOG achieved a total weighted score of 561 and CANLOG planner scored 458. Therefore PLANLOG had an eighteen percent higher weighted score than CANLOG Planner. When contrasted with the sum of the unweighted satisfaction scores this measure emphasised what has been demonstrated in the other sections of the report. With few exceptions, PLANLOG was rated higher than CANLOG Planner.

36. A simple form of sensitivity analysis was conducted on the weight structure to identify how much impact weight variance would have on the evaluation results. In this case, the results were quite in-sensitive to changes. Even when the weights were intentionally skewed to favour CANLOG Planner, PLANLOG retained its overall high score. Other evaluations of software tools may not be this in-sensitive to weight fluctuations

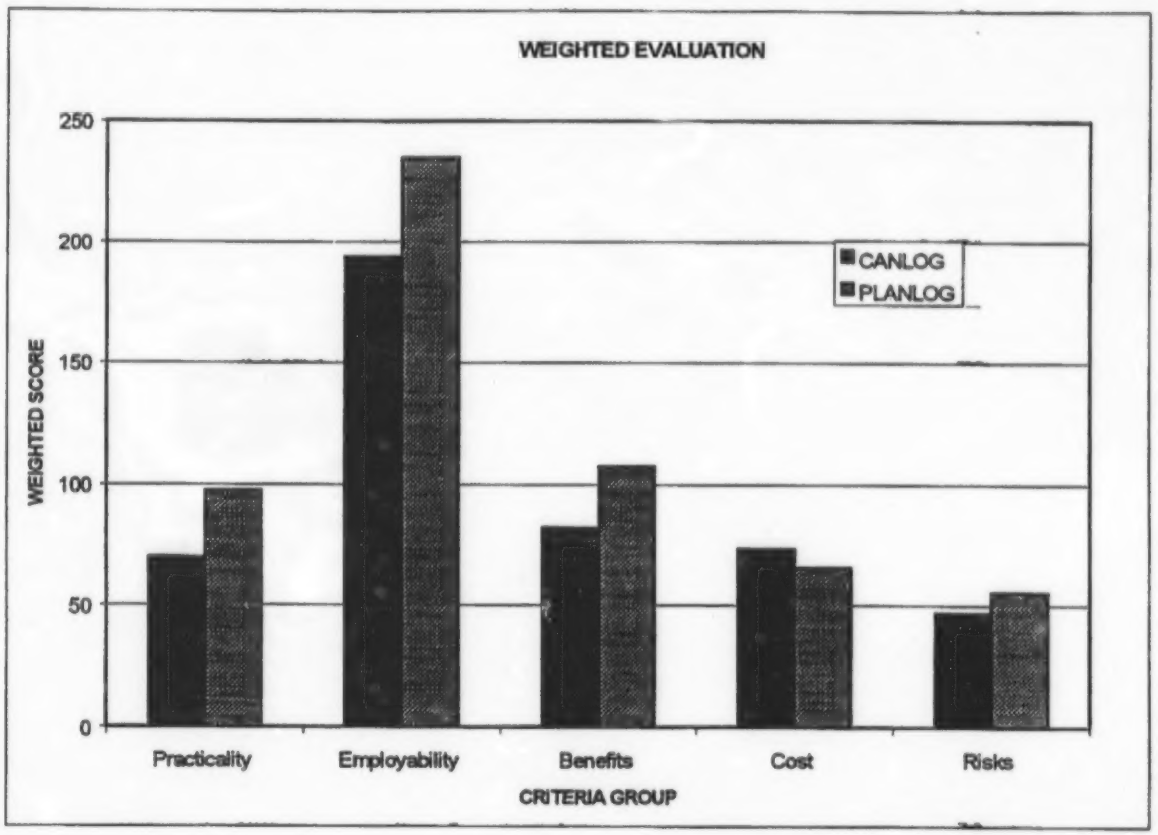

Figure 6: Weighted Evaluation Results

CONCLUSIONS

37. This project has demonstrated that the proposed logistics software evaluation methodology described in this report and reference 3 is very effective. It provides a comprehensive and unbiased evaluation of tools that may not have identical capabilities. It enables a selection to be made between different logistics planning tools. This methodology is reusable. Logistics managers can employ it with limited operations

research assistance. With minor modifications this methodology can be employed for other types of software also.

38. The results of the criteria importance survey offered useful insights into the important issues regarding logistics software. The unanimous emphasis of capabilities over cost issues is an important result.

39. The technical evaluation identified the strengths and weaknesses of both tools. Neither tool had technical weaknesses that precluded them from being employed in DND. Both tools were suitable for inclusion in the user evaluation stage of the evaluation process. Through a direct comparison of the technical evaluation satisfaction ratings, PLANLOG appeared to be technically superior to CANLOG Planner.

40. The user evaluation confirmed some of the findings of the technical evaluation and identified some other concerns. It emphasised the importance of intuitive and user friendly software. The correction of the weaknesses of CANLOG Planner in this area must become a priority if this tool is to remain viable. The employability criteria identified the limitations of both tools. In both cases work needs to be done if the tools are to be used by maritime logisticians. The satisfactory and near-satisfactory scores for PLANLOG in the planning stage criteria emphasised the usefulness of the tool throughout all phases of operations.

41. The evaluation of weighted satisfaction scores confirmed the superiority of PLANLOG over CANLOG Planner. Even when the criteria satisfaction scores were adjusted to accommodate the importance of the individual criterion, PLANLOG rated higher than CANLOG Planner did in four of five criteria groups. With a total weighted score seventeen percent higher than CANLOG Planner's score, it was easy to identify PLANLOG as a superior candidate for use by logisticians in DND

RECOMMENDATIONS

42. As a result of their work on this project, the authors make the following recommendations. The methodology identified in this project report should be employed to evaluate other logistics planning tools under consideration by DND. The use of this methodology will enable a thorough and unbiased consideration of the tools before costly and time-consuming decisions are made.

43. Due to the overwhelming superiority of PLANLOG over CANLOG Planner in this evaluation, it is recommended that the acquisition of PLANLOG for Canadian use be pursued vigorously. This tool offers greater performance and potential to the logistics community without a significant difference in the costs associated with it. Because TNO laboratories have continued to develop PLANLOG during the evaluation period, the newest release of the tool should be examined to determine what changes are required before the tool can be employed by Canada. Once the tool has been acquired for use in DND it should be made available to all environmental staffs and training establishments. This will allow the same tool to be employed for joint or environmentally specific operations and training.

REFERENCES

1. 3451-9(J4 Log). SERVICE PAPER ON MATERIAL/FINANCIAL ACCOUNTABILITY ON DEPLOYED OPERATIONS. 22 May 1996

2. 3451-9(J4 Log Ops) Scaling in Support of Deployed Ops – Statement of Requirement 6 September 1996

3. RN 9803 AN EVALUATION OF LOGISTICS PLANNING TOOLS, Michel Vigneault, Major Doug Hoyle, February 1997

4. 3220-SPT (J4 Log Plans 3) SPONSOR REVIEW – LOGISTICS PLANNING SOFTWARE EVALUATION REPORT 22 Jun 1998

EVALUATION CRITERIA

28 criteria were developed for employment with the evaluation methodology. They have been subdivided into five criteria groups. These groups of criteria have been selected for evaluation by either a technical expert or a logistics planning expert. For quick reference, the criteria have been listed below by criteria group under the type of evaluation they were subjected to.

TECHNICAL EVALUATION

Costs
- Cost of development
- Cost of acquisition
- Cost of Training
- Cost of O&M

Benefits
- Data maintainability
- Growth potential
- Completeness
- Doctrinal fit
- Quality of results
- Supportability

Risks
- Risk of senior management support
- Risk of user acceptance
- Risk of obsolescence

USER EVALUATION

Practicality
- User friendliness
- Learning curve
- Relevance of output
- User-confidence in output

Data sharing

Employability
- Employability in army
- Employability in air force
- Employability in maritime forces
- Employability in joint forces
- Employability for OOTW
- Employability crisis/war
- Employability for what-if scenario analysis
- Employability during pre-deployment
- Employability during deployment
- Employability during employment

CRITERIA DEFINITIONS

Listed below are definitions for all of the evaluation criteria by criteria group.

BENEFIT CRITERIA

Data maintainability refers to the degree to which the data employed by a tool can be maintained with ease. A highly satisfactory level for data maintainability would be that the data could be updated regularly from existing sources, with minimal manipulations. A satisfactory level would be that the data could be updated regularly, with a reasonable investment of effort. A highly unsatisfactory level would be that the data could only be updated regularly, and only through an unreasonable investment of effort.

Growth potential refers to the degree to which the system can be expanded, either by improving the existing functionality, or by extending the functionality of the system. A highly satisfactory level would be that the system could be expanded easily. An unsatisfactory level would be that improving or adding functionality would require a large investment of effort or resources.

Completeness refers to the degree to which the software system performs the full range of logistics planning tasks. The required tasks are not at this point cast in stone, but they generally include determining the logistics requirement, planning for its delivery and tracking the support process. A highly unsatisfactory rating would be given to a software tool that performs no task suitably. A satisfactory rating would be given in cases where the software system performs a broad range of logistics planning tasks.

Doctrinal fit refers to the degree to which the model corresponds to the doctrine of the Canadian Forces. A highly satisfactory rating would be given in the case of a planning tool that accommodates all aspects of logistics doctrine correctly. A satisfactory rating could be given if a minor aspect of logistics doctrine was incorrectly accommodated. An unsatisfactory rating could be given to a tool that accommodates incorrectly an important aspect of logistics doctrine.

Quality of results refers to the degree to which the results are precise and accurate. A highly satisfactory rating would be given to a planning tool for which all results are deemed to be precise and accurate. An unsatisfactory rating would be given to a planning tool for which key results are incorrectly or inaccurately derived.

Supportability refers to the level of support that the developer may be able to give for the software tool in the future. A highly satisfactory rating would be given to a software tool that has a well-established and readily accessible developer. An unsatisfactory rating would be given to a tool for which developer support is likely to be difficult to obtain.

COST CRITERIA

Cost of development refers to the unspent cost of research and development required to build a Canadian version of the software. A satisfactory cost depends on the budget available, which changes over time. As a baseline, we set a nil or negligible cost as a highly satisfactory level, then assess higher costs using the assumption that costs clearly below a reasonable budget level would be satisfactory and costs clearly beyond a reasonable budget level would be unsatisfactory. In the absence of a reasonable budget

level, comparing the expected R&D costs of the different tools can establish satisfaction levels.

Cost of acquisition refers to the cost of purchasing the hardware and software required to support the system. Again, we set a nil or negligible cost as a highly satisfactory level, then assess higher costs relatively to each other, with the assumption that a cost clearly below a reasonable budget level would be satisfactory and a cost clearly beyond a reasonable budget level would be unsatisfactory.

Cost of training refers to the cost of training the personnel required to operate and maintain the system. An evaluation of the training requirements for each software tool will have to be made. A highly satisfactory rating would be given to a tool which uses standards common to other tools and which the users can start operating with little or no training. An unsatisfactory rating would be given to a software tool that requires users to undergo extensive training before being able to operate or maintain the tool. Satisfaction levels can be assessed by comparing the relative training costs for each tool.

Cost of O&M refers to the cost of operating and maintaining the system. A highly satisfactory rating would be given to a tool which requires nil or negligible O&M cost. An unsatisfactory rating would be given to a tool that likely requires an extensive amount of O&M cost. Satisfaction levels can be assessed by comparing the O&M costs relatively to each other.

RISK CRITERIA

Risk of senior management support refers to the risk that senior management will not support the software tool. A highly satisfactory rating would be given to a software tool that has been used in a defence environment and has gained the support of senior management in that environment. An unsatisfactory rating would be given to a software tool that has not been accepted by any senior management or is not likely to be accepted by DND's senior management.

Risk of user acceptance refers to the risk that the users may not accept the software tool. A highly satisfactory rating would be given to a software tool that is known to have a favourable reputation among its current users. An unsatisfactory rating would be given to a software tool that has a significant risk of not being accepted by its users.

Risk of obsolescence refers to the risk that technological advancement could render the system obsolete. A system can be rendered obsolete due to technology changes or when a better system replicates its functionality. A highly satisfactory rating would be given to a software tool that has been developed to accommodate current technology and which has no known competitor likely to render it obsolete. An unsatisfactory rating would be given to a tool that has a significant likelihood of becoming obsolete in the near or medium term.

PRACTICALITY CRITERIA

User friendliness refers to the degree to which the software system incorporates features which make it user-friendly for its users, that is the degree that the software tool responds to users in a way that they feel is pleasant. A highly unsatisfactory rating should be given to a system that includes ongoing irritants that make it difficult to understand and to use. A system, which is easy to understand and to use, should be given a satisfactory rating.

Learning curve refers to the time and effort required to understand the system and to use it productively. A highly satisfactory rating should be given to a system which can be understood and which produces results immediately as it is being used, without the need for initial training. An unsatisfactory rating should be given to a system, which requires a long learning period before it can be understood or used.

Relevance of output refers to degree to which the output is relevant to the logistics planning process. The logistics planning process includes the determination of the logistics requirement, the planning for its delivery and the tracking of the support

process. The complexity of the output, for example, should be appropriate for the task for which the output is being used. A highly unsatisfactory rating should be given to a tool that produces output that is not relevant for the logistics planning process. A satisfactory rating should be given to a software tool that produces output that is mostly relevant and does not omit key relevant output.

User-confidence in output refers to the degree to which the user has confidence in the output. This criterion concerns the way the results are presented: whether the user can track the reasons for a particular output. A highly satisfactory rating means that the results are presented in a way that inspires confidence in the output. An unsatisfactory rating means that the results are not presented in a way that inspires confidence in the output.

Data sharing refers to the degree to which data can be transferred to and from existing logistics systems such as the CFSS, ADAMS, FIS, with minimal difficulty. A highly satisfactory rating means that the software system can export the necessary data files directly into these tools, and that it can directly import the necessary files from these tools. An unsatisfactory rating means that the necessary data cannot be shared electronically with the existing logistics systems.

EMPLOYABILITY CRITERIA

Employability in army refers to the degree to which the software can be useful for logistics planning in the army environment. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning in the army environment. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning in the army environment.

Employability in air force refers to the degree to which the software can be useful for logistics planning in the air force environment. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning in the air force environment. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning in the air force environment.

Employability in maritime forces refers to the degree to which the software can be useful for logistics planning for maritime forces. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning in the maritime environment. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning in the maritime environment.

Employability in joint forces refers to the degree to which the software can be useful for logistics planning for joint force operations. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning for joint forces. A satisfactory rating should be given to a logistics tool which has most of the capability required for logistics planning for joint forces.

Employability for OOTW refers to the degree to which the software can be useful for logistics planning in Operations Other Than War. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning for OOTW. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning for OOTW.

Employability for crisis/war refers to the degree to which the software can be useful for logistics planning for crisis/war. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning for crisis/war. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning for crisis/war.

Employability for what-if scenario analysis refers to the degree to which the software can be useful for logistics planning while performing what-if scenario analysis. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning during the what-if scenario analysis stage. A satisfactory rating should be given to a logistics tool which has most of the capability required for logistics planning during the what-if scenario analysis stage.

Employability during pre-deployment refers to the degree to which the software can be useful for logistics planning during the pre-deployment stage. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning during pre-deployment. A satisfactory rating should be given to a logistics tool which has most of the capability required for logistics planning during pre-deployment.

Employability during deployment refers to the degree to which the software can be useful for logistics planning during the deployment stage. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning during the deployment stage. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning during the deployment stage.

Employability during employment refers to the degree to which the software can be useful for logistics planning during the employment stage. A highly unsatisfactory rating should be given to a logistics tool that has no capability for logistics planning during the employment stage. A satisfactory rating should be given to a logistics tool that has most of the capability required for logistics planning during the employment stage.

EVALUATION MEASUREMENT SYSTEM

The ratings requested from respondents quantified their professional judgement. Importance ratings recorded how much importance stakeholders gave to a criterion while satisfaction ratings recorded how satisfied the evaluators were with the system being considered. Responses could be conveniently recorded by picking a number within a range, with the smallest number representing total lack of satisfaction/importance and the highest number representing the highest level of satisfaction/importance. A scale of one to nine was used because it contained enough gradations to differentiate the potential responses, without overwhelming the respondent. The scale, employed in this evaluation methodology, employed the following levels:

Rating	Definition
1:	**totally unsatisfactory/unimportant**
2:	very unsatisfactory/unimportant
3:	unsatisfactory/unimportant
4:	mildly unsatisfactory/unimportant
5:	**neither unsatisfactory/unimportant nor satisfactory/important**
6:	mildly satisfactory/important
7:	satisfactory/important
8:	very satisfactory/important
9:	**totally satisfactory/important**

ANNEX C
ORD PR 9806
JULY 1998

PLANLOG USER EVALUATION

Section 1 - Evaluation Instructions:

1. PLANLOG is being evaluated for possible employment in DND. PLANLOG has already undergone a technical evaluation. The results of the technical evaluation and the user evaluation will be combined to create an overall rating of PLANLOG. The objective of this evaluation is to gather users input on the strengths and weaknesses of PLANLOG.

2. The user trial aims to examine mainly the functional characteristics of the models. Users will make judgements concerning the level of satisfaction with the functional criteria. A general understanding of the criteria and guidelines for the corresponding satisfaction levels are given in Section 2 of this document. Satisfaction levels have been defined by setting a baseline level at one end of the satisfaction spectrum and adjusting other levels relatively to that baseline. Definitions were chosen to be as objective as possible.

3. Enclosed is a table to record you opinion on PLANLOG. The list of evaluation criteria is shown in the first column on the table, with definitions in the second column. To complete the questionnaire, check **one** box per criteria to indicate your satisfaction with that criteria, using the following definitions:

Rating	Definition
1:	**totally unsatisfactory**
2:	very unsatisfactory
3:	unsatisfactory
4:	mildly unsatisfactory
5:	**neither unsatisfactory nor satisfactory**
6:	mildly satisfactory
7:	satisfactory
8:	very satisfactory
9:	**totally satisfactory**

3. Comments may be included in the last column, using additional pages if necessary. User comments are welcome and strongly encouraged.

4. Questions or queries may be addressed to the POC, Maj. Doug Hoyle, (613) 992-3735, by Banyan e-mail (Maj DGC Hoyle@DGOR@DGOR) or Internet (dhoyle@ora.dnd.ca).

5. Please return the completed questionnaire to the POC by fax (613) 992-3342 or by electronic means no later than **5 November 1997.** An electronic copy of this document has been installed on your PC in the PLANLOG directory.

TABLE I: PLANLOG USER EVALUATION RESULTS

CRITERIA	DEFINITION	IMPORTANCE RATING (Check one box per row) Totally unsatisfactory — totally satisfactory									COMMENTS
		1	2	3	4	5	6	7	8	9	
User friendliness	The degree to which the software system incorporates features which make it user-friendly.										
Learning curve	The time and effort required to understand the system and to use it productively.										
Relevance of output	The degree to which the output is relevant to the logistics planning process.										
User-confidence in output	The degree to which the user has confidence in the output.										
Data sharing	The degree to which data can be transferred to and from existing logistics systems, with minimal difficulty.										
Employability in airforce	The degree to which the software can be useful for logistics planning in the air environment.										
Employability in army	The degree to which the software can be useful for logistics planning in the army environment.										
Employability in maritime forces	The degree to which the software can be useful for logistics planning for maritime forces.										
Employability in joint forces	The degree to which the software can be useful for logistics planning for joint force operations.										
Employability for OOTW	The degree to which the software can be useful for logistics planning in Operations Other Than War.										
Employability for crisis/war	The degree to which the software can be useful for logistics planning for crisis/war operations.										
Employability for what-if scenario analysis	The degree to which the software can be useful for logistics planning while performing "what-if " scenario analysis.										
Employability during pre-deployment	The degree to which the software can be useful for logistics planning during the pre-deployment stage of an operation.										

TABLE I: PLANLOG USER EVALUATION RESULTS											
CRITERIA	DEFINITION	IMPORTANCE RATING (Check one box per row) Totally unsatisfactory totally satisfactory									COMMENTS
		1	2	3	4	5	6	7	8	9	
Employability during deployment	The degree to which the software can be useful for logistics planning during the deployment stage of an operation.										
Employability during employment	The degree to which the software can be useful for logistics planning during the employment stage of an operation.										
Additional comments:											

Section 2 - Evaluation Criteria

The evaluation of PLANLOG requires an examination of a broad range of criteria. The criteria have been divided according to whether they can best be evaluated through a technical evaluation or through a user trail. The breakdown of the criteria between evaluation methods is as follows:

TECHNICAL EVALUATION

Costs
- Cost of development
- Cost of acquisition
- Cost of Training
- Cost of O&M

Benefits
- Data maintainability
- Growth potential
- Completeness
- Doctrinal fit
- Quality of results
- Supportability

Risks
- Risk of senior management support
- Risk of user acceptance
- Risk of obsolescence

USER EVALUATION

Practicality
- User friendliness
- Learning curve
- Relevance of output
- User-confidence in output

Data sharing

Employability
- Employability in army
- Employability in air force
- Employability in maritime forces
- Employability in joint forces
- Employability for OOTW
- Employability crisis/war
- Employability for what-if scenario analysis
- Employability during pre-deployment
- Employability during deployment
- Employability during employment

A general definition of the user evaluation criteria is provided. In all cases at least one end of the range of satisfaction levels is defined to ease the evaluation process.

User friendliness refers to the degree to which the software system incorporates features which make it user-friendly for its users, that is the degree that the software tool responds to users in a way that they feel is pleasant. A highly unsatisfactory rating should be given to a system that includes ongoing irritants that make it difficult to understand and to use. A system which is easy to understand and easy to use should be given a satisfactory rating.

Learning curve refers to the time and effort required to understand the system and to use it productively. A highly satisfactory rating should be given to a system which can be understood and which produces results immediately as it is being used, without the need for

initial training. An unsatisfactory rating should be given to a system that requires a long learning period before it can be understood or used.

Relevance of output refers to degree to which the output is relevant to the logistics planning process. The logistics planning process includes the determination of the logistics requirement, the planning for its delivery and the tracking of the support process. The complexity of the output, for example, should be appropriate for the task for which the output is being used. A highly unsatisfactory rating should be given to a tool that produces output that is not relevant for the logistics planning process. A satisfactory rating should be given to a software tool that produces output that is mostly relevant and does not omit key relevant output.

User-confidence in output refers to the degree to which the user has confidence in the output. This criterion concerns the way the results are presented: whether the user can track the reasons for a particular output. A highly satisfactory rating means that the results are presented in a way that inspires confidence in the output. An unsatisfactory rating means that the results are not presented in a way that inspires confidence in the output.

Data sharing refers to the degree to which data can be transferred to and from existing logistics systems such as the CFSS, ADAMS, FIS, with minimal difficulty. A highly satisfactory rating means that the software system can export the necessary data files directly into these tools, and that it can directly import the necessary files from these tools. An unsatisfactory rating means that the necessary data cannot be shared electronically with the existing logistics systems.

The employability criteria are a means for the users to determine the limits of the logistics-planning tool. There are three employability aspects that are considered: the service (army, air force, maritime, and joint forces), the level of conflict (OOTW, crisis/war) and the planning stage (what-of scenario analysis, pre-deployment, deployment, and employment).

Employability in army refers to the degree to which the functionality of the software can be useful for logistics planning in the army environment. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning in the army environment. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning in the army environment.

Employability in air force refers to the degree to which the functionality of the software can be useful for logistics planning in the air force environment. A highly unsatisfactory rating

should be given to a logistics tool that has none of the functionality required for logistics planning in the air force environment. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning in the air force environment.

Employability in maritime forces refers to the degree to which the functionality of the software can be useful for logistics planning for maritime forces. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning in the maritime environment. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning in the maritime environment.

Employability in joint forces refers to the degree to which the functionality of the software can be useful for logistics planning for joint force operations. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning for joint forces. A satisfactory rating should be given to a logistics tool which has most of the functionality required for logistics planning for joint forces.

Employability for OOTW refers to the degree to which the functionality of the software can be useful for logistics planning in Operations Other Than War. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning for OOTW. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning for OOTW.

Employability for crisis/war refers to the degree to which the functionality of the software can be useful for logistics planning for crisis/war. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning for crisis/war. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning for crisis/war.

Employability for what-if scenario analysis refers to the degree to which the functionality of the software can be useful for logistics planning while performing what-if scenario analysis. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during the what-if scenario analysis stage. A satisfactory rating should be given to a logistics tool which has most of the functionality required for logistics planning during the what-if scenario analysis stage.

Employability during pre-deployment refers to the degree to which the functionality of the software can be useful for logistics planning during the pre-deployment stage. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during pre-deployment. A satisfactory rating should be given to a logistics tool which has most of the functionality required for logistics planning during pre-deployment.

Employability during deployment refers to the degree to which the functionality of the software can be useful for logistics planning during the deployment stage. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during the deployment stage. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning during the deployment stage.

Employability during employment refers to the degree to which the functionality of the software can be useful for logistics planning during the employment stage. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during the employment stage. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning during the employment stage

CANLOG PLANNER USER EVALUATION

Section 1 - Evaluation Instructions:

1. CANLOG PLANNER is being evaluated for possible employment in DND. CANLOG PLANNER has already undergone a technical evaluation. The results of the technical evaluation and the user evaluation will be combined to create an overall rating of CANLOG PLANNER. The objective of this evaluation is to gather users input on the strengths and weaknesses of CANLOG PLANNER.

2. The user evaluation aims to examine mainly the practical characteristics of the model. Users will make judgements concerning the level of satisfaction with several practical criteria. A general understanding of the criteria and guidelines for the corresponding satisfaction levels are given in Section 2 of this document. Satisfaction levels have been defined by setting a baseline level at one end of the satisfaction spectrum and adjusting other levels relatively to that baseline. Definitions were chosen to be as objective as possible.

3. Enclosed is a table to record you opinion on CANLOG PLANNER. The list of evaluation criteria is shown in the first column on the table, with definitions in the second column. To complete the questionnaire, check **one** box per criteria to indicate your satisfaction with that criteria, using the following definitions:

Rating	Definition
1:	**totally unsatisfactory**
2:	very unsatisfactory
3:	unsatisfactory
4:	mildly unsatisfactory
5:	**neither unsatisfactory nor satisfactory**
6:	mildly satisfactory
7:	satisfactory
8:	very satisfactory
9:	**totally satisfactory**

3. Comments may be included in the last column, using additional pages if necessary. User comments are welcome and strongly encouraged.

4. Questions or queries may be addressed to the POC, Maj. Doug Hoyle, (613) 992-3735, by Banyan e-mail (Maj DGC Hoyle@DGOR@DGOR) or Internet (dhoyle@ora.dnd.ca).

5. Please return the completed questionnaire to the POC by fax (613) 992-3342 or by electronic means no later than **9 Feb 1998.**.

TABLE I: CANLOG PLANNER USER EVALUATION RESULTS

CRITERIA	DEFINITION	IMPORTANCE RATING (Check one box per row) Totally unsatisfactory totally satisfactory									COMMENTS
		1	2	3	4	5	6	7	8	9	
User friendliness	The degree to which the software system incorporates features which make it user-friendly.										
Learning curve	The time and effort required to understand the system and to use it productively.										
Relevance of output	The degree to which the output is relevant to the logistics planning process.										
User-confidence in output	The degree to which the user has confidence in the output.										
Data sharing	The degree to which data can be transferred to and from existing logistics systems, with minimal difficulty.										
Employability in airforce	The degree to which the software can be useful for logistics planning in the air environment.										
Employability in army	The degree to which the software can be useful for logistics planning in the army environment.										
Employability in maritime forces	The degree to which the software can be useful for logistics planning for maritime forces.										
Employability in joint forces	The degree to which the software can be useful for logistics planning for joint force operations.										
Employability for OOTW	The degree to which the software can be useful for logistics planning in Operations Other Than War.										
Employability for crisis/war	The degree to which the software can be useful for logistics planning for crisis/war operations.										

TABLE I: CANLOG PLANNER USER EVALUATION RESULTS

CRITERIA	DEFINITION	IMPORTANCE RATING (Check one box per row) Totally unsatisfactory ... totally satisfactory									COMMENTS
		1	2	3	4	5	6	7	8	9	
Employability for what-if scenario analysis	The degree to which the software can be useful for logistics planning while performing "what-if " scenario analysis.										
Employability during pre-deployment	The degree to which the software can be useful for logistics planning during the pre-deployment stage of an operation.										
Employability during deployment	The degree to which the software can be useful for logistics planning during the deployment stage of an operation.										
Employability during employment	The degree to which the software can be useful for logistics planning during the employment stage of an operation.										
Additional comments:											

Section 2 - Evaluation Criteria

The evaluation of CANLOG PLANNER requires an examination of a broad range of criteria. The criteria have been divided according to whether they can best be evaluated through a technical evaluation or through a user trial. The breakdown of the criteria between evaluation methods is as follows.

TECHNICAL EVALUATION

Costs
- Cost of development
- Cost of acquisition
- Cost of Training
- Cost of O&M

Benefits
- Data maintainability
- Growth potential
- Completeness
- Doctrinal fit
- Quality of results
- Supportability

Risks
- Risk of senior management support
- Risk of user acceptance
- Risk of obsolescence

USER EVALUATION

Practicality
- User friendliness
- Learning curve
- Relevance of output
- User-confidence in output

Data sharing

Employability
- Employability in army
- Employability in air force
- Employability in maritime forces
- Employability in joint forces
- Employability for OOTW
- Employability crisis/war
- Employability for what-if scenario analysis
- Employability during pre-deployment
- Employability during deployment
- Employability during employment

A general definition of the user evaluation criteria is provided. In all cases at least one end of the range of satisfaction levels is defined to ease the evaluation process.

User friendliness refers to the degree to which the software system incorporates features which make it user-friendly for its users, that is the degree that the software tool responds to users in a way that they feel is pleasant. A highly unsatisfactory rating should be given to a system that includes ongoing irritants that make it difficult to understand and to use. A system which is easy to understand and easy to use should be given a satisfactory rating.

Learning curve refers to the time and effort required to understand the system and to use it productively. A highly satisfactory rating should be given to a system which can be understood and which produces results immediately as it is being used, without the

need for initial training. An unsatisfactory rating should be given to a system that requires a long learning period before it can be understood or used.

Relevance of output refers to degree to which the output is relevant to the logistics planning process. The logistics planning process includes the determination of the logistics requirement, the planning for its delivery and the tracking of the support process. The complexity of the output, for example, should be appropriate for the task for which the output is being used. A highly unsatisfactory rating should be given to a tool that produces output that is not relevant for the logistics planning process. A satisfactory rating should be given to a software tool that produces output that is mostly relevant and does not omit key relevant output.

User-confidence in output refers to the degree to which the user has confidence in the output. This criterion concerns the way the results are presented: whether the user can track the reasons for a particular output. A highly satisfactory rating means that the results are presented in a way that inspires confidence in the output. An unsatisfactory rating means that the results are not presented in a way that inspires confidence in the output.

Data sharing refers to the degree to which data can be transferred to and from existing logistics systems such as the CFSS, ADAMS, FIS, with minimal difficulty. A highly satisfactory rating means that the software system can export the necessary data files directly into these tools, and that it can directly import the necessary files from these tools. An unsatisfactory rating means that the necessary data cannot be shared electronically with the existing logistics systems.

The employability criteria are a means for the users to determine the limits of the logistics-planning tool. There are three employability aspects that are considered: the service (army, air force, maritime, and joint forces), the level of conflict (OOTW, crisis/war) and the planning stage (what-of scenario analysis, pre-deployment, deployment, and employment).

Employability in army refers to the degree to which the functionality of the software can be useful for logistics planning in the army environment. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning in the army environment. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning in the army environment.

Employability in air force refers to the degree to which the functionality of the software can be useful for logistics planning in the air force environment. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning in the air force environment. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning in the air force environment.

Employability in maritime forces refers to the degree to which the functionality of the software can be useful for logistics planning for maritime forces. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning in the maritime environment. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning in the maritime environment.

Employability in joint forces refers to the degree to which the functionality of the software can be useful for logistics planning for joint force operations. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning for joint forces. A satisfactory rating should be given to a logistics tool which has most of the functionality required for logistics planning for joint forces.

Employability for OOTW refers to the degree to which the functionality of the software can be useful for logistics planning in Operations Other Than War. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning for OOTW. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning for OOTW.

Employability for crisis/war refers to the degree to which the functionality of the software can be useful for logistics planning for crisis/war. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning for crisis/war. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning for crisis/war.

Employability for what-if scenario analysis refers to the degree to which the functionality of the software can be useful for logistics planning while performing what-if scenario analysis. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during the what-if scenario analysis stage. A satisfactory rating should be given to a logistics tool which has most of the functionality required for logistics planning during the what-if scenario analysis stage.

Employability during pre-deployment refers to the degree to which the functionality of the software can be useful for logistics planning during the pre-deployment stage. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during pre-deployment. A satisfactory rating should be given to a logistics tool which has most of the functionality required for logistics planning during pre-deployment.

Employability during deployment refers to the degree to which the functionality of the software can be useful for logistics planning during the deployment stage. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during the deployment stage. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning during the deployment stage.

Employability during employment refers to the degree to which the functionality of the software can be useful for logistics planning during the employment stage. A highly unsatisfactory rating should be given to a logistics tool that has none of the functionality required for logistics planning during the employment stage. A satisfactory rating should be given to a logistics tool that has most of the functionality required for logistics planning during the employment stage.

UNCLASSIFIED

Security Classification of Form
(Highest Classification of Title, Abstract, Keywords)

DOCUMENT CONTROL DATA
(Security classification of title, body of abstract and indexing annotation must be entered whep the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared e.g. Establishment Sponsoring a contractor's report, or tasking agency, are entered in Section 8).

Operational Research Division
Department of National Defence
Ottawa, Ontario K1A 0K2

2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title)

An Evaluation of Logistics Planning Tools

4. AUTHORS (last name, first name, middle initial)

Hoyle, Major Doug G.C. and Vigneault, Michel

5. DATE OF PUBLICATION (month Year of Publication of document)

JULY 1998

6a. NO OF PAGES (total containing information. Include Annexes, Appendices, etc.) 52

6b. NO OF REFS (total cited in document) 4

7. DESCRIPTIVE NOTES (the category of document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Project Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address).

J4 Log

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

002/97 – 3561-LDO-G

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written.)

10a. ORIGINATOR's document number (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

ORD Project Report PR 9808

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

---

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors: further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitations to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11 is possible), a wider announcement audience may be selected.)

—

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

In March 1997, J4 Log tasked LOGORT to conduct an evaluation of two software systems being considered as potential logistics planning tools for the Canadian Forces. A multi-criteria methodology for evaluating logistics planning software was developed. It proposed a list of criteria to be considered. Subject matter experts rated the criteria on their importance to logistics software selection. A two-stage process involving first a technical evaluation and secondly a user evaluation assessed the two tools against the criteria. During each step, a subset of criteria was evaluated using the judgements of evaluators, in order to produce satisfaction ratings. The satisfaction ratings were analysed to identify the strengths and weaknesses of the tools. Subsequently the satisfaction ratings were combined with the importance ratings for the criteria to generate overall performance results. The results of this analysis identified PLANLOG as a better logistics planning tool than CANLOG Planner. This report contains the results of this work and the recommendations that were created from it.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified . If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Logistics Planning
Software Tools
Multi-Criteria Methodology
Decision Support
Technical Evaluation
User Evaluation
Satisfaction Ratings
Importance Ratings
Costs
Benefits
Risks
Practicality
Employability
Planlog
CANLOG Planner
Consumption Estimation
Software Evaluation